

GenCore version 5.1.4 p5 4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 1, 2003, 08:50:46 ; Search time 45.5 Seconds
(without alignment)
5869.482 Million cell updates/sec

Title: US-09-768-781-2

Perfect score: 2543

Sequence: 1 atgaacacagaccaca.....caaggcaagtgtgtctga 1389

Scoring table:

BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09768781/runat_01042003_084730_4952/app_query.fasta_1.1543
-DB=PIR_73 -QMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptp -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09768781@cgn_1_174@runat_01042003_084730_4952 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	932	36.6	444	2	McLeod syndrome-as
2	131	5.2	439	2	hypothetical prote
3	121.5	4.8	783	2	hypothetical prote
4	120.5	4.7	382	2	hypothetical prote
5	120.5	4.7	745	2	cytochrome b245 be
6	116.5	4.6	506	2	sopre germination
7	116.5	4.6	2143	2	hypothetical prote
8	115.5	4.5	498	2	NADH2 dehydrogenas
9	114	4.5	508	2	hypothetical prote
10	112	4.4	459	2	NADH2 dehydrogenas
11	111.5	4.4	785	2	hypothetical prote
12	111	4.4	459	2	NADH2 dehydrogenas
13	111	4.4	459	2	NADH2 dehydrogenas
14	110.5	4.3	459	2	NADH2 dehydrogenas

15	109.5	4.3	461	2	B97305	probable cation ef
16	109.5	4.3	483	2	T20895	hypothetical prote
17	109	4.3	1154	2	T48829	related to SREBP c
18	108.5	4.3	741	2	F90739	probable transport
19	108.5	4.3	741	2	H85589	probable transport
20	108.5	4.3	786	2	H64817	probable membrane
21	108	4.2	740	2	AH0600	probable membrane
22	106.5	4.2	341	2	G82880	ferric anguibactin
23	106.5	4.2	412	2	T34138	hypothetical prote
24	106	4.2	443	2	C70190	K+ transport prote
25	105.5	4.1	339	2	C86874	ABC transporter pe
26	105.5	4.1	503	2	D75104	transmembrane tran
27	105	4.1	459	2	T17181	NADH2 dehydrogenas
28	105	4.1	459	2	T17144	NADH2 dehydrogenas
29	104.5	4.1	493	2	C85655	oligopeptide perme
30	104.5	4.1	493	2	D72059	peptide ABC transp
31	104.5	4.1	521	2	T15322	hypothetical prote
32	103.5	4.1	378	2	T25628	hypothetical prote
33	103.5	4.1	599	2	G90476	probable Na+/H+ an
34	102.5	4.0	460	2	T13881	NADH2 dehydrogenas
35	102	4.0	474	2	S63658	NADH2 dehydrogenas
36	102	4.0	570	2	D86304	hypothetical prote
37	102	4.0	1584	2	T00026	brain-specific ang
38	101.5	4.0	346	2	T19008	hypothetical prote
39	101.5	4.0	354	2	B72688	probable high-affi
40	101.5	4.0	379	2	D96506	hypothetical prote
41	101	4.0	384	2	S37608	NADH2 dehydrogenas
42	100.5	4.0	264	1	S02157	H+-transporting tw
43	100.5	4.0	358	2	B83808	hypothetical prote
44	100.5	4.0	780	2	H84685	probable vacuolar
45	100	3.9	370	2	C69309	conserved hypothet

ALIGNMENTS

RESULT 1

I39294

McLeod syndrome-associated protein XK - human

N:Alternate names: probable membrane transport protein

C:Species: Homo sapiens (man)

C>Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 01-Dec-2000

C:Accession: I39294; S69126

R:Ho, M.; Chelly, J.; Carter, N.; Danek, A.; Crocker, P.; Monaco, A.P.

Cell 77, 869-880, 1994

A:Title: Isolation of the gene for McLeod syndrome that encodes a novel membrane transpo

A:Reference number: A54300; MUID:94273191; PMID:8004674

A:Accession: I39294

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-444 <RES>

A:Cross-references: EMBL:Z32684; NID:G515872; PID:G515873

R:Khamilichi, S.; Bailly, P.; Blanchard, D.; Goossens, D.; Cartron, J.P.; Bertrand, O.

Eur. J. Biochem. 228, 931-934, 1995

A:Title: Purification and partial characterization of the erythrocyte Kx protein defici

A:Reference number: S69126; MUID:95255304; PMID:7737196

A:Accession: S69126

A>Status: preliminary

A:Molecule type: protein

A:Residues: 7-22 <KHA>

C:Genetics:

A:Gene: GDB:XK

A:Cross-references: GDB:120499; OMIM:314850

A:Map position: Xp21.2-Xp21.1

C:Keywords: phosphoprotein; transmembrane protein

Alignment Scores:

Pred. No.: 3.05e-80 Length: 444
Score: 932.00 Matches: 179
Percent Similarity: 64.43% Conservative: 80
Best Local Similarity: 44.53% Mismatches: 133
Query Match: 36.65% Indels: 10
DB: 2 Gaps: 4

US-09-768-781-2 (1-1389) x I39294 (1-444)

```
QY 136 TTTCATTAGCATCCCTTTTCCACCTTTTGTACTGTGGGAGGCTGCATCTGCTTGG 195
Db 3 PheProAlaSerValLeuAlaSerValPheLeuPheValAlaGluThrThrAlaAlaLeu 22
QY 196 TACATGCTTGAATCTATCGAAGAANAATAGTGAACCTTACTCGATGACATACACCTTTCT 255
Db 23 SerLeuSerSerThrTyArgSerGlyGlyAspArgMetTrpGlnAlaLeuThrLeuLeu 42
QY 256 TTCTTTATGTTTTCATCCATTATGGTCAGTTGACCTTATTTTGTCCACAGAGATCTA 315
Db 43 PheSerLeuLeuProCysAlaLeuValGlnLeuThrLeuLeuPheValHisArgAspLeu 62
QY 316 GCCAAGATAAACCGCTATCATTTATGATGATCTAATCTCTTGGGACCTGTTATCAGA 375
Db 63 SerArgAspArgProLeuValLeuLeuLeuHisLeuLeuGlnLeuGlyProLeuPheArg 82
QY 376 TGTTTGGAGGCCATGATTAAGTACTCTCACACTGTGGAGAAAGAGGACGAGGAGGCC 435
Db 83 CysPheGluValPheCysIleTy-----PheGlnSerGlyAsnAsnGluGluPro 99
QY 436 TATGTCAGCTCACCAGAAAGAG---ATGCTAATAGATGGCGAGGAGGTGCTGATAGAA 492
Db 100 TyrValSerIleThrLysLysArgGlnMetProLysAsnGlyLeuSerGluGluIleGlu 119
QY 493 TGGAGGTGGGCCACTCCATCCGACCTGCTATGACCCCAATGCTCAACACGCTATG 552
Db 120 LysGluValGlyGlnAlaGluGlyLysLeuIleThrHisArgSerAlaPheSerArgAla 139
QY 553 TCACAGATCCAAGCTTCTCGGGCTCAGTCCCGACCTGACCTATCAGCTCTATGTGAGC 612
Db 140 SerValIleGlnAlaPheLeuGlySerAlaProGlnLeuThrLeuGlnLeuTyIleSer 159
QY 613 CTGATCTCTGCAGAGGTTCCCTCGGTAGAGTTGTGTAATGGTATTTTCCCTGCTATCT 672
Db 160 ValMetGlnGlnAspValThrValGlyArgSerLeuLeuMetThrIleSerLeuLeuSer 179
QY 673 GTCACCTATGGGCGCACCTTTGGCAATATGTTGGTATCCAGATCAAGTAGACTAC 732
Db 180 IleValTyrGlyAlaLeuArgCysAsnIleLeuAlaIleLysIleLysTyAspIleTyr 199
QY 733 AAGATTCGCTTGGCCACTAGAGTCTCTGTCATCACCTGCGGACATTTGGAGATC 792
Db 200 GluValLysValAsnArgLeuAlaTyValCysIlePheLeuThrPargSerPheGluIle 219
QY 793 ACTTCCGCGCTCTGATTCGTGCTCTTCTCAGCCACTTTGAAATTTGAAGCTGTGCC 852
Db 220 AlaThrArgValValValLeuValLeuPheThrSerValLeuLysThrTrpValVal 239
QY 853 TTCCTAGTGTCAACTTCTGATCATCTCTTTGAGCCCTGGATTAAAGTTCTGGAGAGT 912
Db 240 IleIleLeuIleAsnPhePheSerPhePheLeuTyProTrpIleLeuPheTrpCysSer 259
QY 913 GGTGCCAGATGCCAATAACATTGAGAAATCTCAGCCGGTCGGCAGCTCTGGTGTC 972
Db 260 GlySerProPheProGluAsnIleGluLysAlaLeuSerArgValGlyThrIleVal 279
QY 973 CTGATTTTCAGTCACCATCTCTATCTCGCATCACTTCTCTTGTGCTGACGTTTGAG 1032
Db 280 LeuCysPheLeuThrLeuLeuTyThrGlyIleAsnMetPheCysTrpSerAlaValGln 299
QY 1033 TTGAGTTGGCAGACAGAGATCTCTGACAAAGGCGAAGACTGGGACATATGGCCCTG 1092
Db 300 LeuLysIleAspSerProAspLeuIleSerLysSerHisAsnTrpTyGlnLeuLeuVal 319
QY 1093 CACTATAGTGTGAGTGGTAGAGATGTGATCATGGCTTGGTTTAAAGTTCTTGA 1152
Db 320 TyrTyrMetIleArgPheIleGluAsnAlaIleLeuLeuLeuLeuTyIleLeuPheLys 339
QY 1153 GTGAAAGTGTACTGAATTACTGTCATCTTGTGCTTGTGCTTGCAGTCTATTTGCTAT 1212
Db 340 ThrAspIleTyrMetTyrValCysAlaProLeuLeuValLeuGlnLeuIleGlyTyr 359
```

```
QY 1213 CTGATTTCCATTGGCTTCATGCTCTTTTCTTCAGTACTTGCATCCATTGGCTCACTC 1272
Db 360 CysThrAlaIleLeuPheMetLeuValPheTyGlnPhePheHisProCysLysLeu 379
QY 1273 TTCACCATTAATGTAGTAGAC-----TACCTCCATTGTGTCTGC-----TGT 1314
Db 380 PheSerSerValSerGluGlyPheGlnArgTrpLeuArgCysPheCysTrpAlaCys 399
QY 1315 CACCAG 1320
Db 400 ArgGln 401

RESULT 2
T32470
hypothetical protein F08F1.5 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 04-Mar-2000
C/Accession: T32470
R/Fulton, B.
submitted to the EMBL Data Library, September 1997
A/Description: The sequence of C. elegans cosmid F08F1.
A/Reference number: Z21174
A/Accession: T32470
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-439 <FULL>
A/Cross-references: EMBL:AF026213; PIDN:AAH71305.1; GSPDB:GN000028; CESP:F08F1.5
A/Experimental source: strain Bristol N2; clone F08F1
C/Genetics:
A/Gene: CESP:F08F1.5
A/Map position: X
A/Introns: 20/1; 72/3; 137/3; 184/2; 263/3; 289/3; 339/3; 369/1
C/Superfamily: Caenorhabditis elegans hypothetical protein F08F1.5
```

```
Alignment Scores:
Pred. No.: 0.000618 Length: 439
Score: 131.00 Matches: 67
Percent Similarity: 37.39% Conservative: 62
Best Local Similarity: 19.42% Mismatches: 98
Query Match: 5.15% Indels: 118
DB: 2 Gaps: 15
```

US-09-768-781-2 (1-1389) x T32470 (1-439)

```
QY 328 CCGTATCATTTATTATTCATCTAATCTC-----TTGGGACCTGTTATCATAGTGTG 381
Db 126 ProLeuSerLysMetIleValLeuCysIleCysGlnMetGlyProLeuPheTrpTyr 145
QY 382 GAGGCCATGATTAAGTACCTCACACTGTGGAGAAAGAGGACGAGGAGCCCTATGTC 441
Db 146 LysAlaLeuTy-----TyrGlyTrpMetPheArgLysSerSerAsnGluAsn----- 161
QY 442 AGCCTCACCCGAAAGAGATGCTAATAGATGCGGAGGAGGTGCTGATAGATGGGAGTG 501
Db 162 -----ThrAspGlyGluLys----- 166
QY 502 GGCCACTCATCCCGACCCCTGGCTATGACCGCAATGCCATCAACAGTATGTCACAG----- 558
Db 167 -----ArgLysCysPheSerLysMetValGluAla 176
QY 559 -----ATCCAGAGCTTCTCGGCTCAGTGGCCAGCTG----- 591
Db 177 GluArgAspAlaThrLeuLeuArgPhePheGluAlaPheLeuGluSerAlaProGlnLeu 196
QY 591 ----- 591
Db 197 IleIleGlnGlySerIleAlaAlaSerTyrPheGlnAsnTyrTyGlnThrGlyThrTyr 216
QY 592 ACCTATACGCTCTATGTGAGCCTGATCTCTGCAGAGGTTCCTCGGTAGAGTTGTGCTA 651
Db 217 ProTyrTrpLeuTyPheGlnAlaAlaSer----- 226
```



```
Db 136 leProGluValGluPheSerValValGlyPro-----LysGluAlaPheValG 152
QY 395 AGTACCTCACACTG-----TGGAAGA 415
Db 152 luSerLeuAspLeuAsnIleTyrLeuValAArgArgLeuProThrSerAsnLeuArgI 172
QY 416 AAGAGGAGCAGGAGGAGCCCTATGTCACCTCACCCGAAAGAGATGCTAATAGATGGCG 475
Db 172 leGluGluTyrGluValGlySerLeuSerHisThrArgValAlaIleLeu----- 188
QY 476 AGGAGGTCTGATAGAAAGGGAGGTGGCCCACTCCATCGGACCTGGCTATGACACCGC- 534
Db 189 -----TyrIleGluThrIleThrAsnProGluLeuValThrValValArgAsnArgI 206
QY 535 -----AATGCCACAAACGATGTCACATGTCACAGATCCCAAGCCTCTCCGGCT 577
Db 206 leAsnAspIleAspPheAspHisIleSerAspSerPheIleAsnGlnLeuIleGluA 226
QY 578 CAGTGCCCCAGCTGACCTATCAGCTCTATGTGAGCCTG-----ATCTCTG 622
Db 226 spSerSerGlnSerProPheProLeuLeuIleAspThrGluArgProAspArgValAla 246
QY 623 CAGAGGTTCCTCCCTGGTGTAGTGTGTGTAAGTATTTTCCCTGGTATCTGTACCTATG 682
Db 246 erValLeuAlaGluGlyLysValCysValLeuThrAspGlySerProSerAlaIleThrG 266
QY 683 GGGCCACCTTTGCATATGTTGGTATCCAGATCAAGTACATCACTACAGATTCCGC 742
Db 266 lyProThr-----ThrLeuValGluPhePheSerSerTyrGluAspTyrPheLeuLys 284
QY 743 TTGGGCCACTAGAGTCTCTGCTGCATCACCATCTGCGGACATTTGGAGATCACTCCCGCC 802
Db 284 rpHis-----IleAlaSerIlePheArgLeuIleArgLeuPheSerVal 299
QY 803 TCCTGATTCTG----- 813
Db 299 euPheSerIlePheValThrProMetTyrValAlaValLeuThrTyrHisProGluLeu 319
QY 814 -----GTGCTCTCTCAGCCACTTTGAAATGAGCTGTGCTTC-----CTAG 859
Db 319 leProGlnAspLeuLeuAlaThrLeuThrLeuSerArgSerAlaIleProPheProI 339
QY 860 TGCTCAACTTCTGTATCTCTTTGAGCCCTGATTAAGTTCTGGAGA---AGTGTG 916
Db 339 leLeuGluAlaLeuPheLeu-----GluValThrIleGlnLeuLeuArgGluAlaGly 357
QY 917 CCCAGATGCCCAATAACATTCAGAAAAAATTCAGCGGCTCGGCACCTCTGTGTG----- 969
Db 357 laArgLeuProThrLysValGlyGlnThrIleGlyIleValGlyIleValIleGlyT 377
QY 970 -----GTCTGATTTCAGTCACCATCTCTATG 997
Db 377 hrAlaAlaValGluAlaGlyLeuThrSerAsnValLeuLeuIleIleValAlaLeuAla 397
QY 998 CTGGCATCAACTTCTCTGCTGGTTCAGCTTTCAGCTTTCAGCTTTCAGCTTTCAGCTTTC 1057
Db 397 laLeuAlaSerPheThr----- 403
QY 1058 TCGAACAAGGGCAGAACTGGGACATATGGGCTGCATATAGTGTGAGGTTGGTAGAGA 1117
Db 404 -----ProValTyrArgMetGlyA 410
QY 1118 ATGTGATCATGCTGTGTTTAAAGTTCTTTGGAGTGAAGATGTTACTGAATTAATCTGTC 1177
Db 410 enThrIleArgLeuIleArgPheProPhe-----IleLeuAlaAlaGlnLeu 426
QY 1178 ATTCCTTGATTCCTTCAGCTCATATATGCTTAATCTGATTTCCATTCGCTTCATGCTCC 1237
Db 426 rpGlyLeuIleSerLeuValLeuPheCysTyrLeuLeu-----ValHisLeu 443
QY 1238 TTTTCTTCAG-----TACTTCATCTTCATTCGCTCTCACTCTTCACCCATA 1282
```

```
Db 443 euThrLeuGlnSerLeuGlyArgProPheLeuHisProIleTyrProPheTyrTrpLysA 463
QY 1283 ATGTAGTAGAC 1293
Db 463 spLeuLysAsp 466
RESULT 7
G96595
hypothetical protein F7A10.23 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: G96595
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.; Li, J.H.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: G96595
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-2143 <STO>
A:Cross-references: GB:AE005173; NID:gl0645420; PIDN:AAG21537.1; GSPDB:GN00141
C:Genetics:
A:Gene: F7A10.23
A:Map position: 1
Alignment Scores:
Pred. No.: 0.0198 Length: 2143
Score: 116.50 Matches: 98
Percent Similarity: 33.68% Conservatives: 63
Best Local Similarity: 20.50% Mismatches: 148
Query Match: 4.58% Indels: 169
DB: Gaps: 21
US-09-768-781-2 (1-1389) x G96595 (1-2143)
QY 106 GTCATCCGTGGAGCAACCCCGATTACTTTTCATTTAGCATCTTTTCTCCACTTT 165
Db 678 IleIleAsnSerSerHisGlnPheGluPheGlyPheSerValLeuLeuSerPro 697
QY 166 TTGTACTCTGGGAGGCTGCA-----TCTGCTTTGTACATGGTT 204
Db 698 ValValCysSerIleMetAlaPheLeuArgSerLeuGlnValGluMetAlaLeuThr 717
QY 205 AGAATCTATCGAAAGTAAGTGAACCTTACTGGAGTACATACACC----- 249
Db 718 SerLysSerArgLysTyrGlyPheValAlaTrpLeuLeuSerThrSerValGlyLeuSer 737
QY 250 TTTTCTTTCTTTATCTTTTTCATCCATTTATGTCCTGTCACCTC----- 294
Db 738 LeuSerPheLeuSerLysSerSerValLeuLeuGlyIleSerLeuThrValProLeuMet 757
QY 295 -----ATTTTGTCCACAGA----- 309
Db 758 AlaAlaCysLeuSerIleAlaValProIleTrpMetHisAsnGlyTyrGlnPheTrpVal 777
QY 310 -----GATCTAGCCAAAGAT-----AAACCGCTA 333
Db 778 ProGlnLeuSerCysGlyAspGlnAlaArgAspLeuArgSerProArgIleLysGlyPhe 797
QY 334 TCATTATTATCATCTAATCTCTCTTGGACCTGTTTTCAGATGTTTGGAGGCCATGATT 393
Db 798 IleLeuTrpIleCysValValLeuPheAlaGlySerValIleSerLeuGlyAlaIle 817
QY 394 -----AAGTACCTCACACTGTGGAAGAAAGAGGAGCAG--- 426
Db 818 SerAlaLysProLeuAspAspLeuLysTyrLysLeuPheSerAlaArgGluAsnVal 837
```


NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Pygathrix bieti mitochondrion
C:Species: Mitochondrion Pygathrix bieti
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002
C:Accession: T17163
R:Wang, W.; Foretner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xie, J.
Int. J. Primatol. 18, 305-320, 1997
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.
A:Reference number: Z18709

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-459 <WAN>
A;Cross-references: EMBL:U92956; NID:g2290437; PID:g2290440; PIDN:AAD08923.1
C;Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C;Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:		
Pred. No.:	0.0402	Length: 459
Score:	12.00	Matches: 80
Percent Similarity:	36.54%	Conservative: 72
Best Local Similarity:	19.23%	Mismatches: 142
Query Match:	4.40%	Indels: 122
DB:	2	Gaps: 19

US-09-768-781-2 (1-1389) x T17163 (1-459)

[illegible]

US-09-768-781-2 (1-1389) x A86299 (1-785)

QY 148 ATCCCTTTCTCCACCTTTTGTAC-----TGTGGGAGGCTGCATCT--- 189

RESULT 11

A86299
hypothetical protein F309.18 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
C/Accession: A86299
R/Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chan, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A/Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A86299

A;Status: preliminary

A: Molecule type: DNA

A:Residues: 1-785 <STO>

A: Cross-references: GB:

C:Genetics:

A: Man position: 1

А, маъ рӯстони: 1

Alignment scores:

Alignment scores:

pred. No.:

Score: _____

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-768-781-2 (1-1389

Qy 148 ATCCTTTTCTCCA

Db	30	LeuValPheSerGlnPhePheTyrLeuPheLeuLysProCysGlyGlnAlaGlyProVal	49
Qy	190	-----CCTTGTACATGGTTAGATCTATCGAAGAATAGTGAA	228
Db	50	AlaGlnLeuAlaGlyLeuValLeuSerLeuThrIleLeuArgLeuValHisGlu	69
Qy	229	ACTTACTGGATGACATACACCTTTCTTTCTTATATGTCTTTTCATCATATGCTCCAGCTTG	288
Db	70	PhePheLeuGlnLysAspSerAlaSerTyrTyrIlePhePheSerPheLeuLeuArgThr	89
Qy	289	ACCTCATTTT---GTCCACAGAGATCTAGCCAAAGATAAACCGCTATCATTTATTTATG	345
Db	90	AlaPheValPheLeuIleGlyLeuGluIleAspLeuAsp-----PheMet	104
Qy	346	CATCTAATCTCTTGGGACCTGTTATCATGATGTTTGGAGGCCATGATTAACTACCTCACA	405
Db	105	LysArgAsnLeuLysAsnSerIleValIleThrLeuGlySerLeuValIleSerGlyIle	124
Qy	406	CTGTGGAGAAGAGGAGCAGGAGGAGCCCTATGTCCAGCTCACCCGAAAGAAGATGCTA	465
Db	125	IleTrpLeu-----ProPheLeuTrpPheLeuIleArgPheMetGln	138
Qy	466	ATGATGGCGAG-----	477
Db	139	IleLysGlyAspPheLeuThrPheTyrLeuAlaPheLeuIleThrLeuSerAsnThrAla	158
Qy	478	-----GAGTGTGTATAGATATGGAGGTGGGCCACTGC-----ATCCGAGCC	519
Db	159	AlaProValValIleArgSerIleIleAspTrpLysLeuHisThrSerGluIleGlyArg	178
Qy	520	CTGCGTATGCAACGGCAATCGCTACAAACGTATGTACAGATCCCAAGCCTCTCTGGGCTCA	579
Db	179	LeuAlaIleSerCysGlyLeuPheIleGluIleThrAsnIle-----	192
Qy	580	GTGCCCCAGCTGACCTATCATGCTCTATGTGTAGCGCTGATCTCTGCAGAGGTTCCCTGGGT	639
Db	193	-----PheIleTyrThrIleValLeuSerPheIleSerGlyThrMetThrAlaAsp	209
Qy	640	AGAGTTGTGCTAATGTTATTTCCCTGGTATCTGTCACTATGGGGCCACCCCTTGCAT	699
Db	210	IlePheIleTyrSerPheAlaThrGlyValIleIleLeuThrAsnArgPheLeuAlaSer	229
Qy	700	ATCTTCGCTATCCAGATCAAGTACGATGACTACAAGATTTCGCTTGGGCCACTAGAA---	756
Db	230	TrpLeuProLysArgAsnProLysGluLysTyrLeuSerLysAlaGluThrLeuAlaPhe	249
Qy	757	---GTCTCTGTGCATACCATCTGGCGGACATTGGAG-----ATCAGCTTCCGCGCTC	804
Db	250	IleIleLeuIleLeuIleAlaLeuThrIleGluSerSerAsnLeuAsnSerThrLeu	269
Qy	805	CTGATTCGTGTG-----CTCTTCTCAGCCACTTTGAAA-----TTG	840
Db	270	PheValPheIleGlyLeuMetPheProArgGluGlyLysThrTyrArgThrLeuIle	289
Qy	841	AAGCGTGTGCCCTCTCTAGTGTCTCACTTCCGTGATCATCTCTTTGAGCCCTCGGATTAAG	900
Db	290	GlnArgLeuSerTyrProIleHisGluPheValLeuProValTyrPheGlyTyrIleGly	309
Qy	901	TTCTGGAGAGTGGTCCCGAGATGCCCAATAACATTGAGAAAAAACTTCAGCCGGTCCGC	960
Db	310	PheArgPheSer-----ValAsnSerLeuThrLysArgHisTyrLeuValLeu	325
Qy	961	ACTCTGGTGTCTCTGATTTTCAGTCACCATCTCTATGTGGATCACTTCTCTTCTGCG---	1017
Db	326	GlyMetThrValAlaLeuSerLeuGlyLysLeuLeuGlyValLeuPheAlaCysSer	345
Qy	1018	-----TGGTCAGCTTTTGCAGTTGAGCTTGGCAGACAGAGAT	1053
Db	346	PheLeuLysIleProLysGlnTyrTrpLeuPheLeuSerThrMetLeuSerValLys---	364
Qy	1054	CTCGTCGACAAAGGGCAACTGGGGACATATGGGCGCTCGCATATAGTGTGAGGTGGTA	1113

Df 365 -----GlyHisIleGlyLeu-----ValLeuLeuasp 373

Qy 1114 GAGATGTGATCATGGTCTTGTTTTTAAGTTCTTTGGAGTGAAGTGTACTCAATTAC 1173
||| |||:::|||||::: |||::: |||:::

Df 374 SerAsn-----LeuMetCysLysTrpPheThrProValVal----- 386
||| |||:::|||||::: |||::: |||:::

Qy 1174 TGTCAATCCCTTCGATGCCCTGCAGCTCATTTATCGCTTATCTGATTTCC-----ATTGGC 1227
||| ::| |||:::|||||::: |||::: |||:::

Df 387 --HisAspMetPheValAlaAlaLeuValIleMetThrLeuLeuSerGlyValIleThr 405
||| ::| |||:::|||||::: |||::: |||:::

Qy 1228 TTCATGCTCCCTTTCTTCAG-----TACTTCATCATTCGCTCATCTTCACC 1278
|||:::|||||::: |||::: |||:::

Df 406 SerLeuLeuLeuArgSerGlnGluYsserPheAlaHisIleLysThrSerLeuGluLeu 425
|||:::|||||::: |||::: |||:::

Qy 1279 CATATGTAGTAGATACCTCCATGTGTCTGCTGT 1314
|||:::|||||::: |||::: |||:::

Df 426 PheAspThrThrGluGluLeuArgValLeuThrCys 437
|||:::|||||::: |||::: |||:::

RESULT 12

Tl7166
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Pygathrix bieti mit
C/Species: mitochondrion Pygathrix bieti
C/Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002
C/Accession: Tl7166
R/Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi,
Int. J. Primatol. 18, 305-320, 1997
A>Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.
A/Reference number: Z18709
A/Accession: Tl7166
A>Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-459 <WAN>
A/Cross-references: EMBL:U92957; NID:g2290441; PID:g2290444; PIDN:AAD08826.1
C/Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C/Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;

Alignment Scores:
Pred. No.: 0.0501 Length: 459
Score: 111.00 Matches: 80
Percent Similarity: 36.54% Conservative: 72
Best Local Similarity: 19.23% Mismatches: 142
Query Match: 4.36% Indels: 122
DB: 2 Gaps: 19

US-09-768-781-2 (1-1389) x Tl7166 (1-459)

Qy 142 TTATGATCCTTTCTTCACCTTTTGTACTGTGGGAGGCTGCATCTGCTTTGTACATG 201
||| |||:::|||||::: |||::: |||:::

Df 118 PheTyrlleLeuPheGluThrThrLeulleProThrLeullellelleThrArgTrpgly 137
||| |||:::|||||::: |||::: |||:::

Qy 202 GTTAAATCTCGAAGAATAAGTGAACCTTACTGGATGATACATACCTTTTCTTTCTTT 261
|||:::|||||::: |||::: |||:::

Df 138 AsnGlnAlaGluArgLeuAsnAlaserThyrPheLeuPheThrLeuThr----- 155
|||:::|||||::: |||::: |||:::

Qy 262 ATGTTTTCATTCATGATGTCGACCTCATTTTGTCCAGAGATCTAGCCAAA 321
|||:::|||||::: |||::: |||:::

Df 156 ----GlySerLeuProleuLeulleMetLeuPheValHisAsnAsnLeuGlySer 173
|||:::|||||::: |||::: |||:::

Qy 322 GATAAACCGCTATCATTTATGATCTAACTCTTGGGACCTGTATCAGATGTTTG 381
|||:::|||||::: |||::: |||:::

Df 174 -----LeuAsnIle----- 176
|||:::|||||::: |||::: |||:::

Qy 382 GAGGCCATGATTAAGTACCTCACACTGTGGAAGAGAGGAGGAGGAGCCCTATGTC 441
|||:::|||||::: |||::: |||:::

Df 177 -----ProleuLeu 179
|||:::|||||::: |||::: |||:::

Qy 442 AGCCTCACCCGAAAGAGATGCTAATAGATGGCGAGGAGTGTGATAGATGGGAGTG 501
|||:::|||||::: |||::: |||:::

Df 180 ThrLeuThrAlaGlnIleLeu-----ThrThrThrTrp----- 190
|||:::|||||::: |||::: |||:::

Qy 502 GCCCACTCATCCGACCTCGGTATGACCCGCAATGCCTACAACGATGATCAGATC 561
|||:::|||||::: |||::: |||:::

Df 191 SerHisAsnLeuThrTrpLeuAlaCys-----Met 200
|||:::|||||::: |||::: |||:::

```
QY 562 CAAGCCTTCTCGGCTCAGTGGCCAGTCCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCT 621
Db 201 MetAlaPheMetValLysMetProLeuTyrGlyLeuHisLeuTrpLeuProLysAlaHis 220
QY 622 GCAGAGTTCCCTCG--GGTAGAGTTGCTAATGGTATTTCCCTGATCTGTACACC 678
Db 221 ValGluAlaProIleAlaGlySerMetValLeuAlaValLeuLeuLysLeuGlyGly 240
QY 679 TATGGG-----GCCACCTTTGCAATATGTGTGGCTATCCAGATCAAGTAGCAT 726
Db 241 TyrGlyMetMetArgLeuThrSerIleLeuAsnProLeuThrGluTyrMetAlaTyr--- 259
QY 727 GACTACAAGATTGGCTGGCCACCTAGAGTCTCTGCATCACCATCTGGCGGACATTG 786
Db 260 -----ProPheLeuMetLeuSerLeuTrpGlyMetIle 270
QY 787 GAGTACATCTCCGCCCTCTGATCTGTGGTCTCTTCAGCCACTTTGAAATGAAGCT 846
Db 271 MetThrSerSerThrCysLeuArgGlnThrAspLeuLysSerLeuIleAlaTyrSerSer 290
QY 847 GTGCCCTTCTAGTGTCAACTTCCTGTATCATCTCTTTGAG--CCCTGGATTAACTTC 903
Db 291 ValSerHisMetAlaLeuValIleMetAlaSerLeuIleGlnThrProTrp----- 307
QY 904 TGGAGAAGTGTGGTCCAGATGTCCTTAATCCTCTTGGGACCTGTTATCAGATGTTG 381
Db 174 -----LeuAsnIle----- 176
QY 382 GAGGCCATGATTAAAGTACCTACATGTGGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGG 441
Db 177 -----ProLeuLeu 179
QY 442 AGCTTCACCCGAAAGAGATGCTAATAGATGGCAGGAGGTGCTAGATAGGAGGAGG 501
Db 180 ThrLeuThrAlaGlnLysLeu-----ThrThrThrTrp----- 190
QY 502 GGCACCTCCATCCGACCTGGTGTGTCACCGCAATGTCACCAACGATGTCACAGATC 561
Db 191 SerHisAsnLeuThrTrpLeuAlaCys-----Met 200
QY 562 CAAGCCTTCTCGGCTCAGTGGCCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCT 621
Db 201 MetAlaPheMetValLysMetProLeuTyrGlyLeuHisLeuTrpLeuProLysAlaHis 220
QY 622 GCAGAGTTCCCTCG--GGTAGAGTTGCTAATGGTATTTCCCTGATCTGTACACC 678
Db 221 ValGluAlaProIleAlaGlySerMetValLeuAlaValLeuLeuLysLeuGlyGly 240
QY 679 TATGGG-----GCCACCTTTGCAATATGTGTGGCTATCCAGATCAAGTAGCAT 726
Db 241 TyrGlyMetMetArgLeuThrSerIleLeuAsnProLeuThrGluTyrMetAlaTyr--- 259
QY 727 GACTACAAGATTGGCTGGCCACCTAGAGTCTCTGCATCACCATCTGGCGGACATTG 786
Db 260 -----ProPheLeuMetLeuSerLeuTrpGlyMetIle 270
QY 787 GAGTACATCTCCGCCCTCTGATCTGTGGTCTCTTCAGCCACTTTGAAATGAAGCT 846
Db 271 MetThrSerSerThrCysLeuArgGlnThrAspLeuLysSerLeuIleAlaTyrSerSer 290
QY 847 GTGCCCTTCTAGTGTCAACTTCCTGTATCATCTCTTTGAG--CCCTGGATTAACTTC 903
Db 291 ValSerHisMetAlaLeuValIleMetAlaSerLeuIleGlnThrProTrp----- 307
QY 904 TGGAGAAGTGTGGTCCAGATGTCCTTAATCCTCTTGGGACCTGTTATCAGATGTTG 381
Db 308 -----SerPheThrGlyAla 312
QY 964 CTGGTGGTCTCTGATT-----TCAGTACCACATCTCTATGTGCTGGGATC 1005
Db 313 IleValLeuMetIleAlaHisGlyLeuThrSerSerMetLeuPheCysLeuAlaAsnSer 332
QY 1006 AACTTCTCTTGTGTGCTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTG 1065
Db 332 AenTyrGluArgThrHisSer---ArgIleMetLeuLeuSerArgGlyLeuGlnThrLeu 351
QY 1066 -----GGCAGAACCTGGGGACATATGGCTGCGACTAT 1098
Db 352 LeuProLeuMetAlaPheTrpPheAlaAlaAsnLeuThrAsnLeuAlaLeuProPro 371
QY 1099 AGTGTGAGTTGGTA---GAGAAATGTGATCATGTGCTTGGTTTAAAGTTCTTTGGAGTG 1155
Db 372 ThrIleAsnLeuIleGlyLeuLeuValMetMetThrSerPheSerTrpSerHisVal 391
QY 1156 AAAGTGTACTGAATCTACTGCTATCTCTGATGCTGCTGATGCTGCTGATGCTGCTGCTG 1215
Db 392 ThrIleMetLeuThr-----GlyLeuAsnMetLeuIleThrAlaLeu 405
QY 1216 ATTTCCATTTGGCTTCATGCTCTTTTCTTCCAG-----TACTTGCAT 1257
Db 406 TyrSerLeu---TyrMetLeuValThrThrGlnArgGluThrLeuThrSerHisIleIle 424
QY 1258 CCATTGGCTCACTCTTACC---CATATGTAGTAGACTACTCCAT 1302
Db 425 AsnMetLysProSerPheThrArgGluAsnMetLeuMetPheMetHis 440
```

RESULT 13

T17172

NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Pygathrix bieti mit

C:Species: mitochondrion Pygathrix bieti

C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002

C:Accession: T17172

R:Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi

Int. J. Primatol. 18, 305-320, 1997

A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.

A:Reference number: Z18709

A:Accession: T17172

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-459 <NAN>

A:Cross-references: EMBL:U92959; NID:G2290449; PID:G2290452; PIDN:AND08832.1

C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4

C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation; (

Alignment Scores:
Pred. No.: 0.0501 Length: 459
Score: 111.00 Matches: 80
Percent Similarity: 36.54% Conservative: 72
Best Local Similarity: 19.23% Mismatches: 142
Query Match: 4.36% Indels: 122
DB: 2 Gaps: 19

US-09-768-781-2 (1-1389) x T17172 (1-459)

```
QY 142 TTTPAGCATCTTTTCTCCACCTTTTGTACTGTGGGAGGCTGCATCTGCTTTGTACATG 201
Db 118 PheTyrIleLeuPheGluThrThrLeuIleProThrLeuIleIleThrArgTyrGly 137
QY 202 GTTAGAATCTATCGAAGAATAGTGAACCTTACTGATGATGATACATACACCTTTTCTTT 261
Db 138 AsnGlnAlaGluArgLeuAsnAlaSerThrTyrPheLeuPheTyrThrLeuThr----- 155
QY 262 ATGTTTTCATCCATTATGTCAGTTCAGCTTATTTTGTCCACAGAGATCTAGCCAAA 321
Db 156 -----GlySerLeuProLeuIleMetLeuPheValHisAsnLeuGlySer 173
QY 322 GATAAACCCGCTCATTTATTTATGTCATTAATCCTCTTGGGACCTGTTATCAGATGTTG 381
Db 174 -----LeuAsnIle----- 176
QY 382 GAGGCCATGATTAAAGTACCTACATGTGGAGAAAGAGGAGGAGGAGGAGGAGGAGGAGG 441
Db 177 -----ProLeuLeu 179
QY 442 AGCTTCACCCGAAAGAGATGCTAATAGATGGCAGGAGGTGCTAGATAGGAGGAGG 501
Db 180 ThrLeuThrAlaGlnLysLeu-----ThrThrThrTrp----- 190
QY 502 GGCACCTCCATCCGACCTGGTGTGTCACCGCAATGTCACCAACGATGTCACAGATC 561
Db 191 SerHisAsnLeuThrTrpLeuAlaCys-----Met 200
QY 562 CAAGCCTTCTCGGCTCAGTGGCCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCT 621
Db 201 MetAlaPheMetValLysMetProLeuTyrGlyLeuHisLeuTrpLeuProLysAlaHis 220
QY 622 GCAGAGTTCCCTCG--GGTAGAGTTGCTAATGGTATTTCCCTGATCTGTACACC 678
Db 221 ValGluAlaProIleAlaGlySerMetValLeuAlaValLeuLeuLysLeuGlyGly 240
QY 679 TATGGG-----GCCACCTTTGCAATATGTGTGGCTATCCAGATCAAGTAGCAT 726
Db 241 TyrGlyMetMetArgLeuThrSerIleLeuAsnProLeuThrGluTyrMetAlaTyr--- 259
QY 727 GACTACAAGATTGGCTGGCCACCTAGAGTCTCTGCATCACCATCTGGCGGACATTG 786
Db 260 -----ProPheLeuMetLeuSerLeuTrpGlyMetIle 270
QY 787 GAGTACATCTCCGCCCTCTGATCTGTGGTCTCTTCAGCCACTTTGAAATGAAGCT 846
Db 271 MetThrSerSerThrCysLeuArgGlnThrAspLeuLysSerLeuIleAlaTyrSerSer 290
QY 847 GTGCCCTTCTAGTGTCAACTTCCTGTATCATCTCTTTGAG--CCCTGGATTAACTTC 903
Db 291 ValSerHisMetAlaLeuValIleMetAlaSerLeuIleGlnThrProTrp----- 307
QY 904 TGGAGAAGTGTGGTCCAGATGTCCTTAATCCTCTTGGGACCTGTTATCAGATGTTG 963
Db 308 -----SerPheThrGlyAla 312
QY 964 CTGGTGGTCTCTGATT-----TCAGTACCACATCTCTATGTGCTGGGATC 1005
Db 313 IleValLeuMetIleAlaHisGlyLeuThrSerSerMetLeuPheCysLeuAlaAsnSer 332
QY 1006 AACTTCTCTTGTGTGCTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTGAGTTG 1065
```

```
Db 333 AsnTyrGluArgThrHisSer---ArgIleMetLeuLeuSerArgGlyLeuGlnThrLeu 351
QY 1066 -----GGGCAGAACTGGGGACATATGGGCTGCACATAT 1098
Db 352 LeuProLeuMetAlaPheTrpTrpPheAlaAlaAsnLeuThrAsnLeuAlaLeuProPro 371
QY 1099 AGTGTGAGTTGGTA---GAGAATGTGATCATGTCTTGGTTTAAAGTTCTTTGGAGTG 1155
Db 372 ThrIleAsnLeuIleGlyLeuLeuValMetMetThrSerPheSerTrpSerHisVal 391
QY 1156 AAAGTGTACTGAATTAATCTGTCATCTTGGTTCGCTGAGCTCATTAATGCTTATCTG 1215
Db 392 ThrIleMetLeuThr-----GlyLeuAsnMetLeuIleThrAlaLeu 405
QY 1216 ATTTCCATGGCTTCAGCTCTCTTTCTTCAG-----TACTTGCAAT 1257
Db 406 TyrSerLeu---TyrMetLeuValThrThrGlnArgGlyThrLeuThrSerHisIlelle 424
QY 1258 CCATTGGCTCACTCTTCACC---CATAATGTAGTACACTACCTCCAT 1302
Db 425 AsnMetLysProSerPheThrArgGluAsnMetLeuMetPheMetHis 440

RESULT 14
T17169
NADH2 dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 4 [similarity] - Pygathrix bieti mit
C:Species: mitochondrion Pygathrix bieti
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 03-Jun-2002
C:Accession: T17169
R:Wang, W.; Forstner, M.R.J.; Zhang, Y.P.; Lui, Z.M.; Wei, Y.; Huang, H.Q.; Hu, H.G.; Xi
Int. J. Primatol. 18, 305-320, 1997
A:Title: A phylogeny of Chinese leaf monkeys using mitochondrial ND3-ND4 gene sequences.
A:Reference number: 218709
A:Accession: T17169
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-459 <WAN>
A:Cross-references: EMBL:U92958; NID:G2290445; PID:G2290448; PIDN:RAD08829.1
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 4
C:Keywords: membrane-associated complex; mitochondrion; NAD; oxidative phosphorylation;
Alignment Scores:
Pred. No.: 0.0559 Length: 459
Score: 110.50 Matches: 86
Percent Similarity: 36.45% Conservative: 70
Best Local Similarity: 20.09% Mismatches: 147
Query Match: 4.35% Indels: 126
DB: 2 Gaps: 19

US-09-768-781-2 (1-1389) x T17169 (1-459)
QY 142 TTTAGCATCTTTTCTCCACCTTTTGTACTGTGGGAGGCTGCATCTGCTTTGTACATG 201
Db 118 PheTyrIleLeuPheGluThrThrLeuIleProThrLeuIleIleIleThrArgTrpGly 137
QY 202 GTTAGAATCTATCGAAGAATAGTGAACCTTACTGTGATGACATACACCTTTCTTTCTTT 261
Db 138 AsnGlnAlaGluArgLeuAsnAlaSerThrTyrPheLeuPheTyrThrLeuThr 155
QY 262 ATGTTTTCATCCATTATGGTCCAGTTGACCTTCATTTTGTCCACAGAGATCTAGCCAAA 321
Db 156 -----GlySerLeuProLeuIleMetLeuLeuPheValHisAsnAsnLeuGlySer 173
QY 322 GATAAACCCTATCATTTATTTATGATCTAATCTCTTGGGACCTGTTATCAGATGTTTG 381
Db 174 -----LeuAsnIle----- 176
QY 382 GAGGCCATGATTAAGTACCTCACACTGTGGAAAGAGAGGAGGAGGAGCCCTATGTC 441
Db 177 -----ProLeuLeu 179
QY 442 AGCTCACCACGAAGAAGATGCTAATAGATGCGGAGGAGGTGCTGATAGATGGGAGGTG 501
Db 442 AGCTCACCACGAAGAAGATGCTAATAGATGCGGAGGAGGTGCTGATAGATGGGAGGTG 501
```

```
Db 180 ThrLeuThrAlaGlnLysLeu-----ThrThrThrTrp----- 190
QY 502 GGCACTCCATCCGGACCCCTGGCTATGACCGCAATGCCTACAAACGTATGTCACAGATC 561
Db 191 SerHisAsnLeuThrTrpLeuAlaCys-----Met 200
QY 562 CAAGCTTCTCTGGGCTCAGTGCCCGCAGCTGACCTATCAGCTCTATGTGAGCCTGATCTCT 621
Db 201 MetAlaPheMetValLysMetProLeuTyrGlyLeuHisLeuTrpLeuProLysAlaHis 220
QY 622 GCAGAGTTCCCTCG-----GGTAGAGTTGTGCTAAGTGTATTTCCCTGGTATCTGTCACC 678
Db 221 ValGlnAlaProIleAlaGlySerMetValLeuAlaValLeuLeuLysLeuGlyGly 240
QY 679 TATGGG-----GCCACCTTTGCAATATGTTGGCTATCCAGATCAAGTACGAT 726
Db 241 TyrGlyMetMetArgLeuThrSerIleLeuAsnProLeuThrGluTyrMetAlaTyr--- 259
QY 727 GACTACAAAGATTCCCTTGGGCCACTAGAGTCTCTGTCATCACCATCTGGCGGACATTG 786
Db 260 -----ProPheLeuMetLeuSerLeuTrpGlyMetIle 270
QY 787 GAGATCACTTCCCGCTCTCTGATTCGTGCTCTCTCAGCCACTTTGAAATGAAGGCT 846
Db 271 MetThrSerSerThrCysLeuArgGlnThrAspLeuLysSerLeuIleAlaTyrSerSer 290
QY 847 GTGCCCTTCTAGTGTCTCAACTTCTGATCATCTCTTTGAG---CCCTGGATTAAAGTTC 903
Db 291 ValSerHisMetAlaLeuValIleMetAlaSerLeuIleGlnThrProTrp----- 307
QY 904 TGGAGAGTGTGTGCCAGATGCCCAATAACATTGAGAAAACTTCAGCGGGCTCGGCAC 963
Db 308 -----SerPheThrGlyAla 312
QY 964 CTGTGTGCTCTGATT-----TCAGTCACCATCTCTATGCTGGCATC 1005
Db 313 IleValLeuMetIleAlaHisGlyLeuThrSerSerMetLeuPheCysLeuAlaAsnSer 332
QY 1006 AACTTCTCTGCTGTGCTGAGTTGAGTTGGCAGACAGAGATCTCGTCGACAAA 1065
Db 333 AsnTyrGluArgThrHisSer---ArgIleMetLeuLeuSerArgGlyLeuGlnThrLeu 351
QY 1066 -----GGGCAGAACTGGGCACATATGGGCTGCACATAT 1098
Db 352 LeuProLeuMetAlaPheTrpTrpPheAlaAlaAsnLeuThrAsnLeuAlaLeuProPro 371
QY 1099 AGTGTGAGTTGGTA---GAGAATGTGATCATGTCTTGGTTTAAAGTTCTTTGGAGTG 1155
Db 372 ThrIleAsnLeuIleGlyLeuLeuValMetMetThrSerPheSerTrpSerHisVal 391
QY 1156 AAAGTGTACTGAATTAATCTGTCATCTTGGTTCGCTGAGCTCATTAATGCTTATCTG 1215
Db 392 ThrIleMetLeuThr-----GlyLeuAsnMetLeuIleThrAlaLeu 405
QY 1216 ATTTCCATGGCTTCATGCTCTCTTTCTTCAGTA-----CTTGCACTCCATGGCTC 1268
Db 406 TyrSerLeu---TyrMetLeuValThrThrGlnArgGlyThrLeuThrSerHisIlelle 424
QY 1269 ACTCTTCACCATTAATGATGATAGA-----CTACCTCCATTG 1304
Db 425 AsnMetAsnProSerPheThrArgGluAsnMetLeuMetPheMetHisMetSerProIle 444
QY 1305 TGCTGTGCTGACACGACCCCTCG 1328
Db 445 IleLeuLeuSerLeuAsnProSer 452

RESULT 15
B97305
probable cation efflux pump (multidrug resistance protein) [imported] - Clostridium acet
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: B97305
R:Noelling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
```

; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: B97305

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-461 <KUR>

A:Cross-references: GB:AB001437; PIDN:AAK01229.1; PID:GI5026374; GSPDB:GN00168

A:Experimental source: Clostridium acetobutylicum ATCC824

C:Genetics:

A:Gene: CAC3295

Alignment Scores:

Pred. No.:	0.0697	Length:	461
Score:	109.50	Matches:	81
Percent Similarity:	39.00%	Conservative:	75
Best Local Similarity:	20.25%	Mismatches:	137
Query Match:	4.31%	Indels:	107
DB:	2	Gaps:	19

US-09-768-781-2 (1-1389) x B97305 (1-461)

```
Qy 238 ATGACATACACCTTTCTTCTTTATGTTT---TCATCCATTATGTCAGTTGACCCCTC 294
      ::::: ::::: ::::: ::::: ::::: :::::
Db 107 IISerIISerPheIleIlePheGlnPheLysSerSerIleLeuTyrSerLeuGly--- 125
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 295 ATTTTGTCCAGAGATCTACCAAGATAAA-----CCGCTATCATTA---TTATG 345
      ::::: ::::: ::::: ::::: ::::: :::::
Db 126 -----AlaSerAspAsnThrTyrProMetAlaValGlnTyrIle 138
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 346 CATCTAATCTCTTGGACCTGTTATCAGATGTTTGGAGGCCATGATTAACTACCTCACA 405
      ::::: ::::: ::::: ::::: ::::: :::::
Db 139 SerIleIleLeuPheGlySerIlePheMetCysLeuAlaSerValMetSerTyrIle--- 157
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 406 CTGTGGAGAAAGAGAGGAGGAGGAGGAGCCCTATGTCCAGCTCACCCGAAAGAGATGCTA 465
      ::::: ::::: ::::: ::::: ::::: :::::
Db 158 -----MetValSerLeuGlyLysThrLysThrLeu 167
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 466 ATAGATGGGGAG-----GAGGTCTGATAGATGGGAGTG----- 501
      ::::: ::::: ::::: ::::: ::::: :::::
Db 168 LeuTyrThrAsnIleValGlyValLeuLeuAsnIleIleLeuAsnTyrValLeuValIle 187
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 502 -----GGCCACTCCATCCGACCTGGCTGTATGCACCGCAATGCC 540
      ::::: ::::: ::::: ::::: ::::: :::::
Db 188 GlnLeuHisMetGlyIleArgGlySerGlyIleAlaThrVal----- 201
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 541 TACAAACGTATGTACAGATCCAAAGCCTTCTCGGCTCAGTCCCGCCAGCTGACCTAT--- 597
      ::::: ::::: ::::: ::::: ::::: :::::
Db 202 -----LeuSerGlnLeuAlaAlaPheValValAlaLeuValGlnPheAlaTyrMet 218
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 598 -----CAGCTCTATGTGAGCCTGATCTCTGCAGAG 627
      ::::: ::::: ::::: ::::: ::::: :::::
Db 219 AsnLysLysGlnAsnPheLysPhePheGluAsnIleSerThrAsnIleIleSerGlyAsp 238
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 628 GTTCCCTCGGTAGAGTTGCTAATGTTATTTCTCCCTGGTATCTGTCCACCTATGGGCC 687
      ::::: ::::: ::::: ::::: ::::: :::::
Db 239 Ile---IleArgGluIleValLeuValGlyPheSerThrPheIleIleGluIleAlaAsp 257
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 688 ACCCTTTGCAATATGTTGGCTATCCAGATCAAGTAC-----GATGACTACAAG 735
      ::::: ::::: ::::: ::::: ::::: :::::
Db 258 AlaValIleSerAlaValLeuAsnAsnValLeuTyrAlaGlyGlyAspSerAlaIle 277
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 736 ATTCCGCTTGGGCCACATAGAAGTCTCTGTGATCACCATCTGGCGGACA---TTGGAGATC 792
      ::::: ::::: ::::: ::::: ::::: :::::
Db 278 IleMetLeuGlyValIleThrLysValTyrMetPheMetPheIleThrValIleGlyIle 297
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 793 ACTTCCCGCTCTCTGATTCCTGGTCTCTCTCA----- 825
      ::::: ::::: ::::: ::::: ::::: :::::
Db 298 SerSerGlyMetGlnProIleValGlyTyrAsnPheGlyAlaGlyAsnTyrLysAla 317
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 826 -----GCCACTTTGAAATTGAAGCGCTGTG-----CCCTTCCTAGTGCTC 864
      ::::: ::::: ::::: ::::: ::::: :::::
```

Search completed: April 1, 2003, 08:59:15

Job time : 63.5 secs

```
Db 318 LysAspIleLeuLysPheSerLeuLysThrValIleIleIleSerAlaPheValTyrVal 337
Qy 865 AACTTTCCTCATCATCTCTTTGAGCCCTGGATTAACTTCTGGAGAAAGTGTGCCAGATG 924
      ::::: ::::: ::::: ::::: ::::: :::::
Db 338 GlyPhe---IleIleTyrAlaGlnProLeuIleGlyPhePheLeuLysAspAlaGlnLeu 356
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 925 CCCAATAACATTGAGAAAACCTTCAGCCGGTGGCGACTCTGGTGGTCTCCTGATTCAGTC 984
      ::::: ::::: ::::: ::::: ::::: :::::
Db 357 ValSerLysThrValSerAlaPheArgIleValIleSerMetLeuProLeuLeuGly--- 375
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 985 ACCATCCTCTATGCTGGCATCAACTTCTCTGCTGCTGCTGAGCTTTGCAGTTGAGTTGGCA 1044
      ::::: ::::: ::::: ::::: ::::: :::::
Db 376 ---IleTyrTyrValAlaIleTyrTyr----- 384
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 1045 GACAGAGATCTCGTCGACAAAGGCGCAACTGGGGACATATGGCCCTGCACCTATAGTGTG 1104
      ::::: ::::: ::::: ::::: ::::: :::::
Db 385 -----GlnAlaIleGlyGluAlaArgIleSerPheIleLeu 396
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 1105 AGTTGGTAGAAGATGTGATCATG-----GTCTTGGTTTTTAAGTTCTTT 1149
      ::::: ::::: ::::: ::::: ::::: :::::
Db 397 SerIleTyrArgGluLeuIleMetPheIleProMetAlaValIleLeuPheLysIleVal 416
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 1150 GGAGTCAAA---GTGTTACTGAATTACTGTCTTCTTCTTGCCTTGCCTGCACCTATTATT 1206
      ::::: ::::: ::::: ::::: ::::: :::::
Db 417 GlyIleAsnGlyValPheIleAlaTyrProLeuThrAspIleIleValIleLeuThrSer 436
      ::::: ::::: ::::: ::::: ::::: :::::
Qy 1207 GCTTATCTCATTTCCATTTGGCTTCTATGCTCTCTTCTTCCAGTACTTGCATCCATTCGCG 1266
      ::::: ::::: ::::: ::::: ::::: :::::
Db 437 ValTyrPheIleArgAlaPheLysGluGlnPheAlaGluGluSerIleProLysArg 456
      ::::: ::::: ::::: ::::: ::::: :::::
```